EINDHOVEN UNIVERSITY OF TECHNOLOGY

Testing 3D chips containing through-silicon vias

Citation for published version (APA):
Marinissen, E. J., & Zorian, Y. (2009). Testing 3D chips containing through-silicon vias. In *2009 International Test Conference* [ET1.1] Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/TEST.2009.5355573

DOI:
10.1109/TEST.2009.5355573

Document status and date:
Published: 01/11/2009

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 25. Oct. 2021

Testing 3D Chips Containing Through-Silicon Vias

Erik Jan Marinissen

IMEC vzw
Kapeldreef 75
B-3001 Leuven
Belgium

erik.jan.marinissen@imec.be

Yervant Zorian

Virage Logic
47100 Bayside Parkway
Fremont, CA 94538
United States of America

yervant.zorian@viragelogic.com

Abstract

Today's miniaturization and performance requirements result in the usage of high-density integration and packaging technologies, such as 3D Stacked ICs (3D-SICs) based on Through-Silicon Vias (TSVs). Due to their advanced manufacturing processes and physical access limitations, the complexity and cost associated with testing this type of 3D-SICs are considered major challenges. This Embedded Tutorial provides an overview of the manufacturing steps of TSV-based 3D chips and their associated test challenges. It discusses the necessary flows for wafer-level and package-level tests, the challenges with respect to test contents and wafer-level probe access, and the on-chip DfT infrastructure required for 3D-SICs.

1 Introduction

The semiconductor industry is on an ongoing quest to integrate more functionality into a smaller form factor with increased performance, lower power, and reduced cost. Traditionally, only two-dimensional planes were used for this: through conventional CMOS scaling, multiple IP cores in a single die (System-on-Chip, SoC), multiple dies in a single package (Multi-Chip Package, MCP), and multiple ICs on a Printed Circuit Board (PCB). More recently, also the third, vertical dimension started to become exploited: System-in-Package (SiP), in which multiple naked dies are vertically stacked in a single IC package, and interconnected by means of wire-bonds to the substrate; and Package-on-Package (PoP), in which multiple packaged chips are vertically stacked.

The latest evolution in this list of innovations is the so-called three-dimensional stacked IC (3D-SIC); a single package containing a vertical stack of naked dies which are interconnected by means of Through-Silicon Vias (TSVs) [1–3]. TSVs are conducting nails which stick out of the back-side of a thinned-down die, which allow that die to be vertically interconnected to another die. TSVs are high-density, low-capacity interconnects compared to traditional wire-bonds, and hence allow for much more interconnects between stacked dies, and these interconnects can operate at higher speeds and lower power dissipation. TSV-based 3D technologies enable the creation of a new generation of 'super chips' by opening up new architectural opportunities [4, 5]. Hence, they allow the semiconductor industry to continue to still its hunger for more functionality, bandwidth, and performance at smaller sizes, power dissipation [6], and cost; even in an era in which conventional feature-size scaling becomes increasingly difficult and expensive.

Like all ICs, also these new TSV-based 3D-SICs need to be tested for manufacturing defects, in order to guarantee sufficient outgoing product quality to the customer. In this Embedded Tutorial paper, we describe the test needs of 3D-SICs, review to what extent existing test solutions can be used to address those, and discuss what new test challenges come with this new class of products.

The remainder of this paper is organized as follows. Section 2 positions TSV-based 3D-SICs in the market place and describes various TSV technology options. Test flows for 2D and 3D ICs are compared in Section 3. Section 4 argues that modular testing is a very appropriate approach for 3D-SICs and lists the required infrastructure for it. Section 5 discusses the new test contents we need for intra-die defects due to new 3D processing steps and for the TSV interconnects. Sections 6 and 7 focus on test access, resp. for wafer test and within the chip through an on-chip DfT architecture. Section 8 concludes this paper.

2 3D-SICs Based on TSV Technology

Micro-electronic systems typically consist of multiple integrated circuits (ICs). Traditionally, these ICs were brought together as a system by interconnecting them on a PCB. Ongoing integration has brought us SoC, MCP, SiP, and PoP, all with their respective benefits and drawbacks [7]. SoCs integrate all functionality in a single die, with the benefit of fast and abundant inter-module communication; however, SoC integration forces all modules to make use of the same (for some design modules non-optimal) process technology and might lead to large, low-yielding 'monster' chips. MCPs, SiPs, and PoPs integrate multiple dies, and hence provide advantages over SoCs with respect to heterogeneous system integration, as these dies can be manu-

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

factured in different, optimized process technologies. All three integration approaches are based on wiring interconnects, which are limited in number and relatively slow and power-consuming. Only SiP and PoP make use of the vertical dimension, and hence offer the benefit of a smaller footprint.

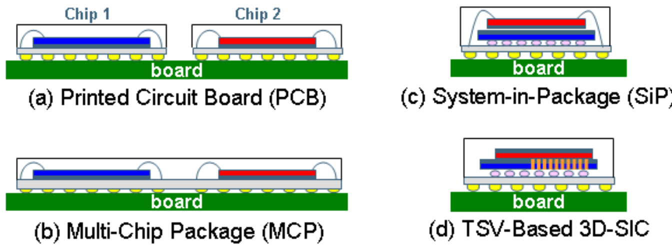


Figure 1: Ongoing system integration evolution: 2D (a) PCB, (b) MCP, and 3D (c) SiP, and (d) 3D-SIC.

A TSV-based 3D-SIC combines the advantages of the above integration technologies [3]. It integrates multiple dies, with as potential benefits smaller, high-yielding dies in possibly different, dedicated process technologies [8]. It utilizes the vertical dimension for stacking thinned dies, thereby creating a smaller footprint as well as a higher transistor density per volume unit. And the TSV-based interconnects between the various dies have in number, speed, and power dissipation more similarity to on-chip wires than to off-chip wire bonds [5, 9]. On-chip interconnect wire length, especially of the global and semi-global wires, can be reduced drastically [10].

The process technology for 3D-SICs based on TSVs is becoming available now, design support tools are emerging [11], and first products start to appear on the market. Among the early applications are CMOS image sensors stacked on their corresponding read-out and digital processing circuitry, and stacks of memories [12]. Memory-on-logic [13] and logic-on-logic [14] are expected to follow suit.

Through-Silicon Vias provide, as their name indicates, an electrical connection from the active front-side ('face') of a silicon die through the silicon substrate to the 'back'-side. This allows to interconnect multiple vertically stacked dies with each other. Below, we describe one of the TSV types developed at IMEC. These TSVs are cylindrical copper nails of 25μm height, 5μm diameter (i.e., an aspect ratio of 5:1), and have a minimum pitch of 10μm [15, 16]. The TSV fabrication steps are depicted in Figure 2 and consist of (1) deep silicon etching of TSV holes, (2) oxide deposition, (3) copper seed deposition, (4) copper plating, and (5) chemical-mechanical polishing (CMP).

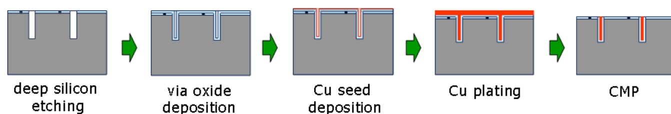


Figure 2: Subsequent TSV fabrication steps.

As can be seen in Figure 2, when the wafer processing is completed, the TSVs are still deeply buried within the wafer; their height is only 25μm, while the total wafer thickness is around 750μm. To expose the TSV tips, the wafer needs to be thinned down from the back-side to just below 25μm thickness. In order to provide sufficient mechanical strength and prevent it from breaking or cracking, the to-be-thinned wafer is temporary bonded onto a carrier wafer, prior to thinning [17]. Subsequently, the thinned product wafer on its carrier wafer is permanently bonded to the next die, after which the temporary carrier wafer is removed. The thinning and bonding steps are depicted in Figure 3. This process can be repeated in case more than two dies are stacked.

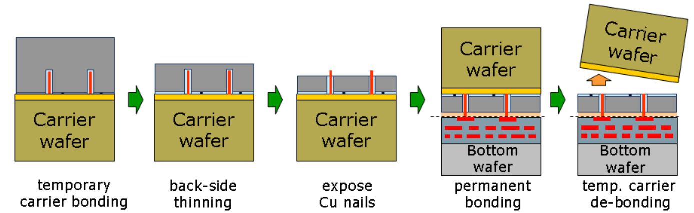


Figure 3: Subsequent wafer thinning and bonding steps.

There are many different variants of TSV types, process steps, and stacking options [3, 18, 19]. Below, we describe some of them, in as far as relevant for the remainder of this paper.

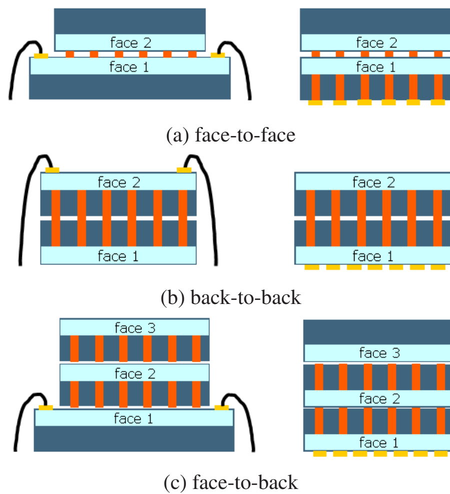
The TSVs we described above have a 5μm diameter, 25μm height, and 10μm minimum pitch; TSVs can also be larger or smaller. Taller TSVs cause less problems with respect to wafer thinning, but have a larger aspect ratio and hence are more difficult to properly fill [20]. That is, unless we also make the TSV diameter larger, but that has a negative impact on the silicon area they consume as keep-out area, the maximum TSV density, and the capacitive load (and performance and power dissipation) of the TSV [7].

TSVs can be fabricated as a 'via-first', 'via-middle', or 'via-last' process. The TSVs described above are fabricated as a 'via-middle' process, i.e., *after* the (Front-End-Of-Line, FEOL) transistors are fabricated, but *before* the (Back-End-Of-Line, BEOL) metal layers [16]. 'Via-first' TSVs are fabricated *before* FEOL, while 'via-last' TSVs are fabricated *after* BEOL. Typically, 'via-first' and 'via-middle' are quite a lot smaller, denser, and with larger aspect ratios than the rather large 'via-last' TSVs. One could say that 'via-first/middle' TSVs more follow the characteristics of semiconductor interconnect technology, while 'via-last' TSVs are closer to assembly interconnect technology.

The orientation of the individual dies in the stack is another item of variation. Dies can be connected 'face-to-face', 'back-to-back', or 'face-to-back' [14], as depicted in Figure 4. 'Face' refers to the active front-side of the die, whereas 'back' refers to the substrate back-side, where the exposed TSV tips stick out. In 'face-to-face' stacking (Figure 4(a)), the active sides of the two dies are interconnected directly. In case the external com-

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

munication is through wire-bonds, no TSVs are necessary; but the bottom die should be slightly larger than the top die to allow space for wire-bonding. In case of flip-chip bumps, TSVs are required make the connection through those. 'Face-to-face' stacking does not scale well to stacks of more than two dies. In 'back-to-back' stacking (Figure 4(b)), TSVs in both dies are used to interconnect the two dies, while the external communication goes via either via wire-bonds from the top die or flip-chip bumps at the bottom die. Also this scenario does not scale well to stacks of more than two dies, and moreover does not provide the minimum amount of TSV-equipped layers (which is important, as the fabrication of TSV comes with extra processing steps and associated costs). 'Face-to-back' stacking (Figure 4(c)) has the advantage that it is scalable to stacks of more than two dies. The TSV tips sticking out of the back of one die connect to dedicated TSV landing pads on the face of the other die. 'Face-to-back' stacking can be used with all dies in the stack face-up, while the external communication is done via wire-bonding from the bottom die; again, it is required that the bottom die is somewhat larger than the other dies. Alternatively, 'face-to-back' stacking also works with all dies in the stack face-down, while the external communication is done via flip-chip bonds at the face-down bottom die.



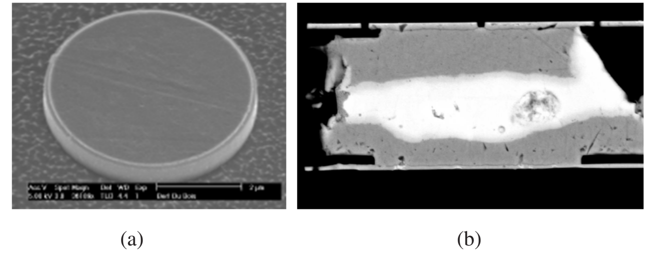
(a) face-to-face

(b) back-to-back

(c) face-to-back

Figure 4: Different stacking orientation variants with wire-bond and flip-chip external connections.

The bonding technology is yet another variation item. The TSVs with 10μm minimal pitch described above have been realized with copper-to-copper (Cu-Cu) direct bonding. The TSV tips are cylindric copper nails with a 5μm diameter (as shown in Figure 5(a)), and the TSV landing pads are copper rectangles of 9×9μm^2. We use thermo-compression bonding at relatively high temperatures. This bonding technology achieves sub-micron stand-off heights between the two dies, which also means that the bond quality is very sensitive to small particles in between both dies [21]. IMEC also have experimented with bonds based on copper-tin (Cu-Sn) micro-bumps between both dies, as depicted in Figure 5(b) [8]. While enlarging the minimum pitch from 10μm to around 50μm (and hence reducing the maximum TSV density), this bonding type works at much lower temperatures and due to the larger stand-off height between the two dies is far less sensitive to particles in between the dies.



(a)

(b)

Figure 5: Photos of two alternative bonding technologies: (a) 5μm-diameter Cu TSV tip of direct Cu-Cu bond process, and (b) inter-metallic bond of a Cu-Sn-Cu micro-bump process.

We distinguish between Wafer-to-Wafer (W2W), Die-to-Wafer (D2W), and Die-to-Die (D2D) stacking. W2W stacking has the benefit that it avoids the pick-and-place operations required for D2W and D2D stacking; however, W2W stacking has far less possibilities to exploit Known-Good Die (KGD) test results, and hence might suffer from low compound yields and corresponding higher costs.

The technology options described in this section (and others we did not discuss) make that there is a wide variety of TSV technologies under consideration right now. Some will prove ineffective or inefficient and will be abandoned; nevertheless, it seems likely that multiple alternative TSV technologies will continue to exist alongside each other [2].

3 3D Test Flow

Conventional single-die chips have two natural test moments: (1) *wafer test* (also referred to as 'e-sort') takes place after wafer fabrication and before assembly and packaging, and (2) *final test* takes place after assembly and packaging. The outgoing product quality is guaranteed by the final test. The wafer test typically serves merely as an economical optimization to prevent unnecessary package and packaging costs of dies that are found to be bad already during wafer test. Hence, whether or not wafer testing should be done for a particular product, depends on the wafer fabrication yield y, the fraction d of faulty products that the test can detect (which is determined by the test quality), the costs p that can be prevented per product (here: the cost of packaging a single die), and the cost t of executing a wafer test on a single product. For many conventional ICs, wafer testing pays off, i.e., its benefits exceed its costs:

$$(1 - y) \cdot d \cdot p > t. \tag{3.1}$$

For multi-die ICs, there are more intermediate product stages, and hence there are more natural test moments [22]. For a 3D-IC consisting of n dies, we have potentially $2n$ natural test moments, as depicted in Figure 6:

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

- n Known-Good Die (KGD) tests of the individual dies before using them;

- $n-2$ Known-Good Stack (KGS) tests of intermediate stacking results (i.e., Dies 1+2, Dies 1+2+3, up to Dies 1+2+...+$n-1$);

- one Known-Good Stack (KGS) test of the final stack result (i.e., Dies 1+2+...+n);

- one final test of the packaged product.

The only test of a packaged product in this list is the final test, which again serves the purpose of guaranteeing the outgoing product quality. All other tests are tests of unpackaged dies and die stacks. As handling and testing individual dies is rather cumbersome, we assume here that all these tests are wafer tests.

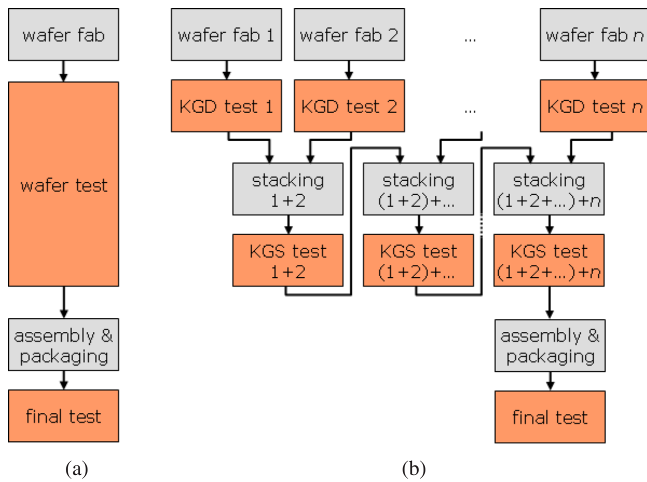


Figure 6: 2D (a) and 3D (b) test flow.

With so many more wafer test opportunities, it seems likely that the importance of wafer testing will increase. Whether or not all these tests make indeed technical and economical sense, is governed by the same trade-off expressed by Equation 3.1.

For 3D-SICs, the costs p that can be prevented per product are typically more comprehensive; it includes all downstream material and processing costs: other dies from the stack, stacking costs, the package, and packaging costs. However, parameter p is also influenced by whether or not we are able to act upon the test outcome. For example, in a W2W stacking approach, we typically cannot avoid to stack an individual bad die. In this case, KGD/KGS test results can be utilized to crank up the stack yield only if we are able to stock wafers in a repository and stack them based on wafer-map *matching*, which seems a rather expensive and cumbersome process.

The parameters y, d, and t are strongly influenced by which faults are targeted during a particular test. For KGD tests, an important question is whether or not we are targeting TSV faults as well; this increases the number $(1-y)\cdot d$ of faulty parts we might catch, but at the same time increases the test execution costs t. For KGS tests, an important question is whether we only test the newly-made latest bond, or also re-test the previously tested dies and bonds in the stack. Re-testing comes with extra test execution costs t, and hence makes sense only if the latest stacking operation indeed can introduce a significant number of new defects in the already tested items.

4 Modular Testing

Modular testing refers to an approach in which the various modules that together constitute an IC product are tested as separate, stand-alone units [23]. For complex SoCs, modular testing is gaining increasing traction, for the following reasons.

- Heterogeneous modules contain different circuit structures and exhibit different defects; consequently, they require dedicated fault models and test patterns. For example: embedded memories and mixed-signal blocks are tested separately from the embedding digital logic.

- Black-boxed IP cores can only be tested with test patterns as delivered by their core providers, as only they know the implementation details of those cores.

- A divide-n-conquer approach in which a monolithic design is broken up into more digestible chunks makes test pattern generation more tractable, and significantly reduces the resulting test data volume as well, as every module gets only the test patterns it requires, instead of the SoC-wide maximum [24].

- Modular testing allows for easy reuse of tests, throughout the life-cycle of one SoC, and also in subsequent derivate designs that reuse the same design module.

The modular test approach is especially suited to 3D-SICs. All four arguments listed above for modular SoC testing apply even stronger to the case of 3D-SICs. In addition, a modular test approach fits very naturally with the KGD/KGS test flow as depicted in Figure 6(b). A modular test approach allows to freely decide where in the test flow a certain module is tested and/or re-tested. Finally, in case of a faulty product, a modular test approach allows a first-order diagnosis, sufficient to pin-point the faulty component (which might be crucial if the overall 3D-SIC product is assembled from components of different sources).

Obvious test modules are the various individual dies, the TSV-based interconnect between dies, and the *'extra-connect'* to the external world. For example, in a two-die stack, we would consider at least four separate test modules: (1) Die 1, (2) Die 2, (3) the interconnect between Dies 1 and 2, and (4) the extra-connect to the product pins. If the dies themselves are complex, possibly heterogeneous SoCs, the modular test hierarchy can be further extended by dividing them into various sub-modules.

Modular testing requires the following infrastructure.

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

1. A language in which the test description is transferred from one party to the next (standardized as IEEE Std. 1450.6, Core Test Language (CTL) [25]).

2. On-chip DfT in the form of a core-test wrapper (standardized as IEEE Std. 1500 [26]) and a Test Access Mechanism (TAM) such as a TestRail or test bus [27].

3. EDA support for automated test expansion, i.e., the translation from module-level test into chip-level test [28].

Section 7 discusses the on-chip DfT to support modular testing, as part of the overall DfT architecture for 3D-SICs, in more detail.

5 3D Test Contents

The test contents of 3D-SICs is, in a first-order approximation, not very different from conventional 2D ICs. In the (FEOL/BEOL) wafer manufacturing, the same defects can occur, and hence the faults, fault models and test patterns are the same. However, in this section, we focus on the differences in test content between conventional 2D ICs and new 3D-SICs. First of all, the 3D processing might lead to new intra-die defects and faults, that require new tests. Secondly, the TSV-based interconnects constitute a new structure, for which we have to define new tests.

5.1 New Intra-Die Defects

Commonly used tests for digital logic, memories, analog, and RF modules cover the majority of known defects. However, for 3D-SICs, the question is whether some of the new 3D processing steps cause new defects that are not covered by conventional fault models.

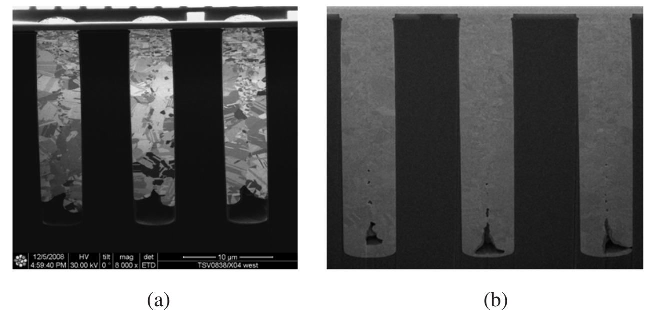
Wafer thinning is such a 3D processing step that can cause new defects. The TSV processing only allows for limited TSV heights (say, 10 to 100μm) and aspect ratios (say, 10:1 maximum) [20]. In order to expose the TSV tips at the back-side for bonding to another die, the wafer needs to be thinned down. Although this is an area for further research, early results indicate degradation of some *I-V* characteristics, shifts in device performances, and limited yield losses [29, 30].

Thermal dissipation and thermo-mechanical stress are other causes of concern. Integrated circuits heat up during operation. In densely packed stacks of thinned dies, the heat density might pile up quite high, and has little way of escaping. The heat generated might easily impact the correct operation of the various dies, especially since some dies are more heat-sensitive (e.g., DRAMs) than others. Due to the different Coefficients of Thermal Expansion (CTEs) of the various materials in the stack, the stack might also suffer from thermo-mechanical stress, causing further malfunction.

5.2 TSV Interconnect Test

The TSV-based interconnect of 3D-SICs constitute a new structure, not existing in conventional 2D ICs. Hence, we need to consider which defects it might suffer from, how these defects behave as faults, how to model these faults, and how to test for them.

TSV-related defects might occur either in the fabrication of the TSV themselves (see Figure 7), in the bonding of the TSVs to the next layer, or during the life time of the 3D stack. During the fabrication of TSVs, (micro-)voids, for example due to quasi-conformal plating, might lead to (weak) opens in TSVs, while ineffective removal of the seed layer might lead to shorts between TSVs. The bond quality might be negatively impacted by oxidation or contamination of the bond surface, height variation of the TSVs, or particles in between the two dies. Misalignment during bonding, in either x, y, or (tilted) z direction might lead to opens or shorts. In case of Cu-Sn micro-bumps, the tin might squeeze out due to TSV height variation and cause shorts in between them. During the product's life time, CTE mismatch might cause the thinned dies to warp, after processing and/or during operation; also, thinned dies might be more susceptible to the effects of mechanical load.



(a) (b)

Figure 7: Examples of TSV defects: (a) insufficiently filled TSVs, and (b) TSVs containing micro-voids along the axis and at the bottom.

Although the actual defect mechanisms are different, the resulting faults for TSV-based interconnects are rather similar to the faults normally considered for wiring interconnects: opens, shorts, and delay faults. This means there is also a large body of existing work with respect to test pattern generation [31] which we can leverage, such as the Counting Sequence Algorithm [32], the Modified Counting Sequence Algorithm [33], and the True/Complement Test Algorithm [34]. These algorithms detect all (hard) open and shorts through a set of digital test patterns that can be kept small, as it grows only logarithmically with the number of interconnects. The test algorithms rely on the fact that we have full controllability at all interconnect inputs and full observability at all interconnect outputs (as is for example the case with Boundary-Scan-Test (IEEE Std. 1149.1) equipped ICs at a PCB). In Section 7 of this paper we discuss this requirement in more detail. Testing delay faults on TSV-based interconnects is not difficult nor test-time consuming either, but requires good

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

synchronization between both dies.

The tests described above all rely on the power-, ground-, and clocking infrastructure in the dies across both sides of the TSV-based interconnect to be present and working. All dies in a 3D-SIC stack, apart from the bottom die, receive their power, ground, and clock signals also through TSVs. Typically, multiple redundant TSVs are used to transfer power and ground, and hence that infrastructure might be less sensitive for faults in single TSVs. On the other hand, this redundancy also complicates the detection in power and ground TSVs. Faults in clock-signal TSVs are typically easier to detect, as a non-transferred clock signal often leads to catastrophic results.

6 3D Wafer Test Access

As argued in Section 3, a 3D test flow contains potentially many more wafer tests for carrying out KGD and KGS tests on intermediate product stages, which prevent downstream processing of bad products and hence are necessary to keep the overall production cost down. Wafer testing of 3D intermediate products has two major challenges connected to it: probing on very many, very small probe points, and the handling of thinned wafers and stacks containing thinned wafers.

Today's probe technology, using either cantilever or vertical probes, goes down to a minimum pitch of 35μm [35], has a maximum probe count of several hundreds, and makes significant scrub marks in order to achieve a proper electrical contact. This is insufficient to probe on TSV tips of 5μm diameter and 10μm pitch (or smaller), which might come in several thousands (the 10μm pitch allows TSV densities up to 10k/mm^2), are made of fragile copper, and do not tolerate scrub marks that inhibit downstream Cu-Cu bonding on the same surface. Probing on Cu-Sn micro-bumps is still challenging, but nevertheless a bit easier, as the sizes and pitches of the micro-bumps are larger, consequently their numbers smaller, and the constraints on scrub marks less strict.

For KGD testing of the various dies that will make up a 3D-SIC, we distinguish between the bottom die and the other (non-bottom) dies. The non-bottom dies receive all their functional signals (power, ground, clocks, control, data) exclusively through TSV connections, and hence only possess I/Os which cannot be probed with today's probe technology. The bottom die is different, in the sense that, next to its TSV connections, it also has wire-bond or flip-chip pads for the extra-connect, which provide an interface probe-able with today's probe technology and which allows us to get test data in and out of that die and test it. If we want to execute KGD tests on the other, non-bottom dies, new solutions need to be developed. The following solution approaches can be explored.

- *Additional probe pads*
  Providing dedicated additional probe pads (as a form of DfT) at the side to be probed and sized such that today's probe technology can handle them. Obviously, this comes with an area penalty, and hence the number of extra pads should be minimized; on-chip DfT can help with this (see Section 7.

- *Probe technology improvement*
  Significant improvement of wafer probe technology, scaling down to (in the order of) 25μm pitch for micro-bump probing or down to (in the order of) 10μm pitch for TSV tip and TSV landing pad probing, while at the same time increasing the maximum probe count and reducing the scrub mark damage.

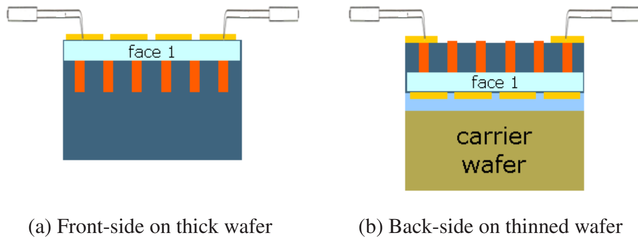
- *Contactless wafer probing*
  Further development of contactless wafer probe technology, for example based on inductive coupling [36, 37]. This technology has the inherent advantage to not cause any scrub marks. However, also this technology needs to scale down in order to probe on micro-bumps and/or TSVs in the sizes and densities they occur. Moreover, the Tx/Rx circuit in the wafer probe card needs to be a mirror image of the DUT and hence might require standardization, while DUT power delivery is *not* contactless and hence still requires conventional contact probes.

While the latter two approaches still require further development effort, the first approach is feasible today.

As an example, let us consider the 3D-SIC depicted in Figure 4(c). It consists of three dies, which are stacked face-down, and face-to-back. The bottom die and middle die contain TSVs, the top die does not. The bottom die provides the extra-connect by means of flip-chip bumps. Figures 8–10 show the various KGD wafer probe options that exist for these three dies.

The bottom die can be probed on its (regular-sized) flip-chip pads on the front-side, as shown in Figure 8(a). The wafer is not yet thinned, and hence its TSVs are buried within the thick wafer and cannot be accessed. Alternatively, it would be attractive to execute the KGD tests on a thinned wafer, in order to be able to also detect defects due to wafer thinning. A thinned wafer only exists on a carrier wafer, to prevent cracking. The flip-chip pads are now inaccessible, due to the carrier wafer. Wafer probe access could be obtained from the back-side of the thinned wafer, where the exposed TSV tips or micro-bumps are located. Unfortunately, these are too small and numerous to be probed by today's probe technology. A solution could be to provide a number of dedicated additional regular-sized probe pads on the back-side (as a form of DfT). This is depicted in Figure 8(b), where the small TSV tips/pads are shown in orange and the larger dedicate probe-able pads in yellow.

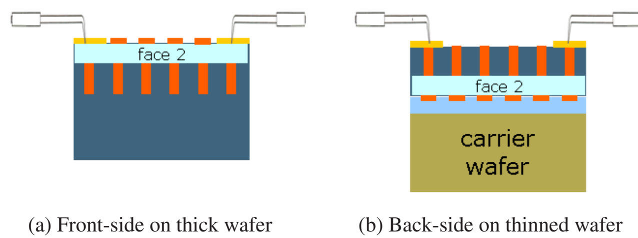
978-1-4244-4867-8/09/$25.00 ©2009 IEEE



(a) Front-side on thick wafer (b) Back-side on thinned wafer

Figure 8: Wafer probe options for the bottom die.

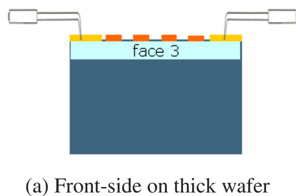
Figure 9 shows the two similar alternatives (front-side probing on a thick wafer and back-side probing on a thinned wafer on carrier) for the middle die. The difference between bottom die and middle die is that the middle die has functional probe-able pads on neither front- nor back-side. Hence, also for the front-side probing (Figure 9(a)), dedicated additional regular-sized probe pads need to be provided.



(a) Front-side on thick wafer (b) Back-side on thinned wafer

Figure 9: Wafer probe options for the middle die.

In our example stacking scenario, the top die does not contain TSVs, and is not thinned. Hence, probing on its back-side is not an option. As shown in Figure 10, this leaves us only with the possibility to probe on the front-side of the thick wafer, where again dedicated additional regular-sized probe pads need to be provided, as the functional TSV landing pads are too small for today's probe technology.



(a) Front-side on thick wafer

Figure 10: Wafer probe options for the top die.

A question that arises is whether the costs associated to providing additional probe pads to be able to execute a KGD test do not entirely consume the benefits of that same KGD test. Using IMEC's elaborate 3D cost model (which accounts for the cost of clean room, equipment, maintenance, materials, and personnel, as well as production yields and test coverage), we have done cost-benefit calculations, which demonstrate that for most manufacturing yields, the costs of providing additional probe pads on the non-bottom dies to enable KGD testing, *does* pay off [38]. As an example, one such trade-off calculation assumed the following case.

- A two-die stack in a D2W or D2D stacking process (such that the KGD tests can be utilized, and bad dies are excluded from stacking).
- Both dies are full-scan testable digital logic dies. The top die contains 15M gates (10×10 mm^2), while the bottom die contains 20M gates (12×12 mm^2).
- The bottom die always undergoes a KGD test. It contains 100 scan chains, and its KGD test uses 2k patterns to achieve 99% test coverage.
- The trade-off is about the KGD test for the top die. If tested, it uses 13 scan chains, and 2k patterns to achieve 99% test coverage. To be able to do this, it requires 30 additional probe pads (with associated silicon area costs). Alternatively, the top die is not tested; we do not need additional probe pads, but are required to blindly stack both good and bad top dies.
- All bad stacks (due to bad top die, bottom die, or TSV interconnect) are detected by the final test. The total manufacturing and test cost can only be recuperated by increasing the cost price of good products.

The results of the cost model calculations are depicted in Figure 11. The horizontal axis shows the wafer manufacturing yield for both top and bottom die, which is assumed to be equal and in order to take into account the different die sizes expressed in yield per cm^2. The vertical axis shows the relative production costs per produced good 3D stack, normalized to the (cheapest) case in which wafer yield is 100% and *no* KGD test for the top die is performed. Figure 11 shows that performing the KGD test pays off for most wafer yields, despite the extra test execution costs and the extra costs for providing extra, probe-able sized pads on the top die.

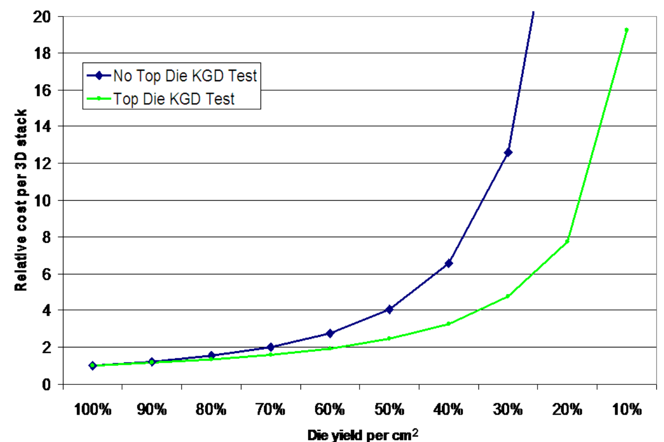


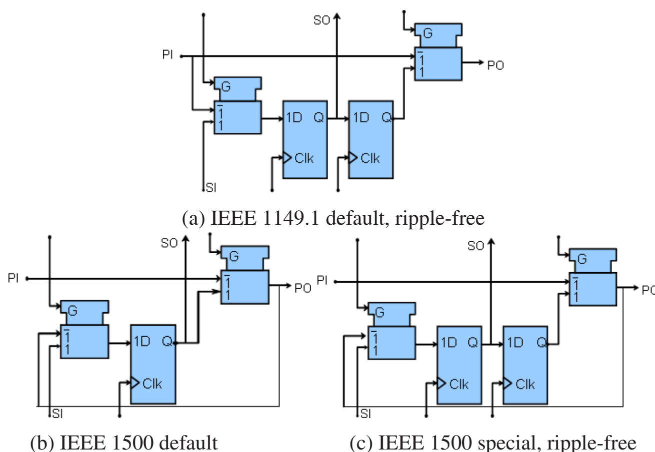
Figure 11: Cost comparison between executing a top-die KGD test or not, where the first alternative requires extra probe pads.

7 3D On-Chip DfT Architecture

The primary role of on-chip Design-for-Testability (DfT) is to provide controllability and observability from the chip I/Os into the heart of the chip design and vice versa. In the basic 3D DfT architecture described below, we assume a 3D-SIC consisting of multiple stacked dies, which each consist of one or multiple 'test modules' (see Section 4). We distinguish DfT at various levels.

1. *Within the test modules*
   DfT within the test modules, e.g., internal scan chains.

2. *Between die-level and module-level*
   DfT to enable modular testing, i.e., wrappers around the test modules and TAMs [23]. Wrappers provide controllability and observability at the module boundary, for both inward-testing (*InTest*) as well as outward-testing (*ExTest*). Wrappers for embedded test modules should be based on IEEE Std. 1500 [26], as this standard was dedicatedly developed for this purpose. Wrappers at the boundary of a die could either be based on IEEE Std. 1149.1 or IEEE Std. 1500 (see Figure 12). Especially if dies are designed by different teams or companies, and not necessarily 'friends' at their interfaces, it is probably a wise thing to equip them with ripple-protected wrapper cells. The TAM widths should be scaled in order to optimize test bandwidth vs. test length [27].



(a) IEEE 1149.1 default, ripple-free

(b) IEEE 1500 default

(c) IEEE 1500 special, ripple-free

Figure 12: Various wrapper cell implementations.

3. *Between (intermediate) stack-level and die-level*
   DfT from the product's ATE interface to the die-level TAMs and vice versa. This includes:

   - For the non-bottom dies: dedicated additional probe pads for wafer test access, (see Section 6), and a switch between test access from the extra probe pads into the local TAM (for KGD tests) and test access from the TSVs of the next-lower die into the local TAM (for KGS and final tests).
   - For the non-top dies: TAMs to transparently pass test stimuli to the TSVs of the next-higher die and to pass test responses from the TSVs of the next-higher die. Also, the DfT in the die should be able to operate independently from the presence (or not) of the die(s) above it.

4. *At the completed stack-level*
   DfT to support board-level testing, once the product is soldered to a Printed Circuit Board (PCB). Towards this end, the overall product should be compatible to IEEE Std. 1149.1/4/6.

For illustration purposes, Figure 13 shows a simplified example of a DfT architecture that meets the above requirements. It consists of two dies, stacked face-down in a face-to-back fashion. Each die is a single scan-testable logic module. The functional I/Os are shown in yellow, while the color light-blue is used for additional DfT: wrapper cells, multiplexers, TSV interconnects, and external pads. The various DfT multiplexer settings allow for five different test access modes, which all (except for the board-test mode) have two scan chains: one for the wrapper boundary register, and one for the module-internal scan chains.

- A mode in which a KGD test can be applied *before* wafer thinning to the bottom die from its front side. Test access is via two scan chains, which are multiplexed onto already-existing functional pads, viz. `D1/SI1–Q1/SO1` and `D2/SI2–Q2/SO2`. Note that in this mode, the wrapper cells at the extra-connect terminals of the bottom die do *not* need to be included in scan chains, as they are directly controllable and observable from the wafer probes.

- A mode in which a KGD test can be applied *after* wafer thinning to the bottom die from its back side. Test access is via two scan chains which connect to dedicated backside pads, viz. `SIb1–SOb1` and `SIb2–SOb2`. Note that in this mode, the wrapper cells at the extra-connect terminals of the bottom die do need to be included in the wrapper boundary register scan chain.

- A mode in which a KGD test can be applied to the top die from its front side. Test access is via two scan chains which connect to dedicated front-side pads, viz. `SIb1–SOb1` and `SIb2–SOb2`.

- A mode in which a KGS test or final test can be applied to the stack of bottom and top die. Test access is via two scan chains which are multiplexed onto already-existing functional pads, viz. `D1/SI1–Q1/SO1` and `D2/SI2–Q2/SO2`. This mode actually consists of three sub-modes, in which respectively the bottom die, the top die, and the TSV-based interconnects between bottom and top die can be tested.

- A mode in which a board-level test is executed. Test access is via one boundary scan chain along the extra-connect interface of the bottom die which connects to dedicated front-

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

side pads, viz. `TDI–TDO` (part of the IEEE 1149.1 Boundary Scan standard).

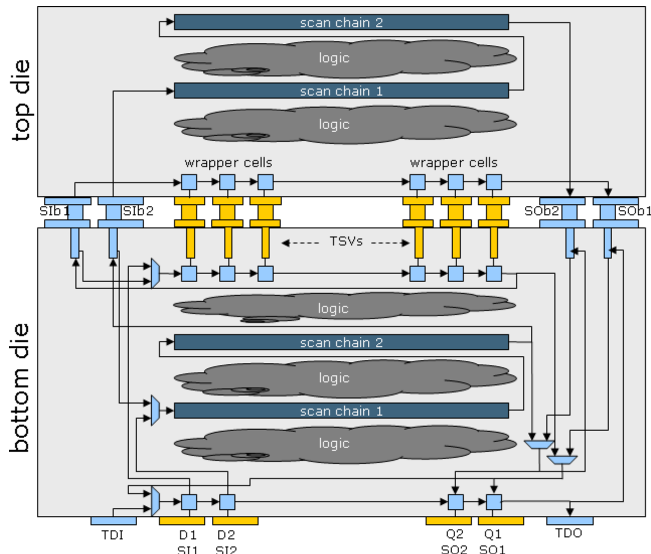


Figure 13: Simplified example of a 3D DfT architecture.

The above example is simplified in various ways. Figure 7 shows only test data, and abstracts from test control signals. The stack consists of two dies only. Each die consists of of a single test module only. Each test module has only two scan chains: one wrapper boundary scan chain and one consisting of two concatenated internal scan chains. In reality, many of these parameters will be (significantly) larger. As is the case of DfT architectures for conventional 2D SoCs [27], it is worthwhile for large real-life 3D-SICs to optimize the DfT architecture, such that the corresponding test schedule allows for a minimal test length [39–42].

The following DfT techniques play a special role in the context of 3D-SICs.

- *Reduced Pad-Count Testing (RPCT)*
  RPCT is a DfT technique to reduce the width of a scan test interface [43]. As discussed in Section 6, dedicated additional probe pads might need to be provided to execute a KGD test on the bottom die *after* thinning or to execute a KGD test on other, non-bottom dies. These extra probe pads are costly in area. RPCT can be exploited to reduce the number of extra pads required. Note that the usage of RPCT does not affect the total test data volume, which implies that a reduced test interface width comes at the expense of an increased test length.

- *Test Data Compression (TDC)*
  TDC is a DfT technique that exploits the many 'don't care' bits in ATPG-generated test patterns to compress the off-chip test data in a (near-) lossless manner. It allows to reduce the volume of both test stimuli and responses (and the corresponding test length) with one or two orders of magnitude [44]. TDC can play a role in reducing the test data volumes of 3D-SICs. When applied per test module, the TAMs that provide these test modules with stimuli and responses can be scaled down. Also, RPCT and TDC form an attractive DfT combination, as the first scales down the test interface width, while the second prevents the test length from increasing.

- *Built-In Self-Test (BIST)*
  BIST is a DfT technique that performs the entire generation of test stimuli and evaluation test responses on chip, such that the IC test becomes completely self-contained and does not require external test equipment (other than to start the BIST and read out its result). BIST can play in role in reducing the test data volumes of 3D-SIC. Actually, BIST reduces the test data volume even more than TDC, viz. down to (virtually) zero. The price to be paid for this is that LBIST (for digital logic) typically requires more implementation area, the test time can be longer, and/or the test quality less. MBIST for (embedded) memories, on the other hand, does not have these drawbacks and seems very attractive for 3D-SICs containing either memory dies or memories embedded in logic dies. Other benefits of BIST are that it might play a role in protecting proprietary test contents, such that a die user can *execute* a test without having to know its exact contents, and in re-using tests across the IC's life time [45, 46].

Just as 3D-SICs offer the opportunity to system architects to re-design and re-optimize their system architecture, these die stacks also offer the opportunity to DfT architects to re-design and re-optimize their DfT architecture. Especially where this involves the decision which DfT resource to put into which die, we refer to this as *Test Resource Partitioning*. Let us illustrate the possibilities by means of an example. Consider a 3D-SIC product consisting of a memory die stacked on top of a logic die. The memory die provider might be used to make only stand-alone memories, which often do not come with BIST; being stand-alone, the memory I/Os are fully accessible from the test equipment, and the test costs are typically reduced by testing many memory chips in parallel ('multi-site'). In a 3D-SIC context, for the KGD test the memory is still a stand-alone memory, but for the KGS test that same memory die is actually more similar to an embedded memory. And, for embedded memories, BIST is the DfT methodology of choice. One can envision a scenario in which the memory die comes '3D-prepared', i.e., with an on-chip BIST engine. This scenario has as advantage that any proprietary memory test content does not need to be released. However, the MBIST (logic) might be difficult to implement in the technology of the memory die (perhaps DRAM). Therefore, an alternative scenario could be that the memory maker provides a 'drop-in' description of a MBIST engine, which is actually implemented in the bottom logic die. The MBIST operation is controlled from within the logic die, whereas the stimuli and re-

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

sponses flow over TSV-based interconnect into and out of the memory die.

8 Conclusion

3D-SICs have many compelling benefits and hence are quickly gaining ground. It is to be expected that soon they will take a significant share of the overall semiconductor market. Test solutions need to be ready for this new generation of 'super chips'. 3D-SICs are chips where all basic as well as most advanced test technologies come together. In addition, they pose some truly new test challenges.

3D-SICs offer many more natural moments to perform wafer tests compared to conventional 2D chips. Performing all these tests will lead to an explosion of test costs, while the main purpose of wafer tests is to reduce overall costs. Finding a test flow with an optimal set of KGD and KGS tests will be a different trade-off for every 3D-SIC product. A modular test approach treats the various components of the product as separate test units and hence fits very well to 3D-SICs. It allows to optimize the test flow, by freely deciding where in that flow a certain module is tested and/or re-tested.

Fault models and corresponding tests for new intra-die defects need to be developed. Examples of causes of such new defects are novel 3D processing steps such as wafer thinning, as well as thermal and thermo-mechanical stress. Fault models and corresponding tests also need to be developed for the TSV interconnects between dies.

Wafer test for 3D-SICs is a challenge, as today's probe solutions do not allow to probe on the I/Os (in the form of TSV tips or micro-bumps) of the non-bottom dies. A way out is to provide extra probeable-sized pads, but a more cost-effective solution can hopefully be found in a significant improvement of wafer probe technologies.

As in all ICs, DfT plays a crucial role in transporting test stimuli and responses in and out of the various test modules. Our paper showed an example DfT architecture that enables modular testing and supports the various KGD/KGS test modes. This architecture can be further augmented with DfT techniques like RPCT, TDC, and BIST. 3D-SICs open the possibility to re-evaluate existing test resource partitions and explore new ones.

Note that this paper only focused on electrical testing of 3D-SICs. There are also many challenges for 3D-SICs in related areas as design validation, in-line inspection, metrology, diagnosis, failure analysis, redundancy and repair, but they are beyond the scope of this paper.

Acknowledgements

We thank many colleagues at IMEC for stimulating discussions, especially Eric Beyne, Luc Dupas, Pol Marchal, Nikolaos Minas, Dan Perry, Geert Van der Plas, Bart Swinnen, and Dimitrios Velenis.

References

[1] Robert S. Patti. Three-Dimensional Integrated Circuits and the Future of System-on-Chip Designs. *Proceedings of the IEEE*, 94(6):1214–1224, June 2006.

[2] Eric Beyne and Bart Swinnen. 3D System Integration Technologies. In *Proceedings of IEEE International Conference on Integrated Circuit Design and Technology (ICICDT)*, pages 1–3, June 2007.

[3] Philip Garrou, Christopher Bower, and Peter Ramm, editors. *Handbook of 3D Integration – Technology and Applications of 3D Integrated Circuits*. Wiley-VCH, Weinheim, Germany, August 2008.

[4] Gabriel H. Loh, Yuan Xie, and Bryan Black. Processor Design in 3D Die-Stacking Technologies. *IEEE Micro*, 27(3):31–48, May/June 2007.

[5] Roshan Weerasekera, Li-Rong Zheng, Dinesh Pamanuwa, and Hannu Tenhunen. Extending Systems-on-Chip to the Third Dimension: Performance, Cost and Technological Tradeoffs. In *Proceedings International Conference on Computer-Aided Design (ICCAD)*, pages 212–219, November 2007.

[6] James W. Joyner and James D. Meindl. Opportunities for Reduced Power Dissipation Using Three-Dimensional Integration. In *Proceedings IEEE International Interconnect Technology Conference (IITC)*, pages 148–150, June 2002.

[7] Bart Swinnen. 3D Technologies: Requiring More Than 3 Dimensions from Concept to Product. In *Proceedings IEEE International Interconnect Technology Conference (IITC)*, pages 59–62, June 2009.

[8] Eric Beyne et al. Through-Silicon Via and Die Stacking Technologies for Micro Systems Integration. In *Proceedings IEEE International Electron Devices Meeting (IEDM)*, pages 1–4, December 2008.

[9] Kaustav Banerjee et al. 3-D ICs: A Novel Chip Design for Improving Deep-Submicrometer Interconnect Performance and Systems-on-Chip Integration. *Proceedings of the IEEE*, 89(5):602–633, May 2001.

[10] Shamik Das, Anantha Chandrakasan, and Rafael Reif. Three-Dimensional Integrated Circuits: Performance, Design Methodology, and CAD Tools. In *Proceedings of IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, pages 13–18, February 2003.

[11] Shamik Das, Anantha Chandrakasan, and Rafael Reif. Design Tools for 3-D Integrated Circuits. In *Proceedings IEEE Asia South Pacific Design Automation Conference (ASP-DAC)*, pages 53–56, January 2003.

[12] Uksong Kang et al. 8Gb 3D DDR3 DRAM Using Through-Silicon-Via Technology. In *Proceedings International Solid State Circuits Conference (ISSCC)*, pages 130–132, February 2009.

[13] Christianto C. Liu et al. Bridging the Processor-Memory Performance Gap with 3D IC Technology. *IEEE Design & Test of Computers*, 22(6):556–564, November/December 2005.

[14] Bryan Black et al. 3D Processing Technology and its Impact on iA32 Microprocessors. In *Proceedings International Conference on Computer Design (ICCD)*, pages 316–318, October 2004.

[15] Bart Swinnen et al. 3D Integration by Cu-Cu Thermo-Compression Bonding of Extremely Thinned Bulk-Si Die Containing 10μm Pitch Through-Si Vias. In *Proceedings IEEE International Electron Devices Meeting (IEDM)*, pages 1–4, May

978-1-4244-4867-8/09/$25.00 ©2009 IEEE

2006.
[16] Jan Van Olmen et al. 3D Stacked IC Demonstration using a Through Silicon Via First Approach. In *Proceedings IEEE International Electron Devices Meeting (IEDM)*, pages 1–4, December 2008.
[17] Alain Jourdain et al. Electrically Yielding Collective Hybrid Bonding for 3D Stacking of ICs. In *Electronic Components and Technology Conference (ECTC)*, pages 11–13, May 2009.
[18] Rafael Reif et al. Fabrication Technologies for Three-Dimensional Integrated Circuits. In *Proceedings International Symposium on Quality of Electronic Design (ISQED)*, pages 33–37, March 2002.
[19] J.U. Knickerbocker et al. 3D Silicon Integration. In *Electronic Components and Technology Conference (ECTC)*, pages 538–543, May 2008.
[20] Bioh Kim et al. Factors Affecting Copper Filling Process Within High Aspect Ratio Deep Vias for 3D Chip Stacking. In *Electronic Components and Technology Conference (ECTC)*, pages 1–6, 2006.
[21] Twan Bearda et al. Post-Dicing Particle Control for 3D Stacked IC Integration Flows. In *Electronic Components and Technology Conference (ECTC)*, pages 1513–1516, May 2009.
[22] Yervant Zorian, editor. *Multi-Chip Module Test Strategies*. Kluwer Academic Publishers, 1997.
[23] Yervant Zorian, Erik Jan Marinissen, and Sujit Dey. Testing Embedded-Core Based System Chips. In *Proceedings IEEE International Test Conference (ITC)*, pages 130–143, October 1998.
[24] Ozgur Sinanoglu et al. Test Data Volume Comparison: Monolithic vs. Modular SoC Testing. *IEEE Design & Test of Computers*, 26(3):25–37, May/June 2009.
[25] Rohit Kapur. *CTL for Test Information of Digital ICs*. Kluwer Academic Publishers, 2003.
[26] Francisco da Silva, Teresa McLaurin, and Tom Waayers. *The Core Test Wrapper Handbook: Rationale and Application of IEEE Std. 1500*. Springer-Verlag, 2006.
[27] Sandeep K. Goel and Erik Jan Marinissen. SOC Test Architecture Design for Efficient Utilization of Test Bandwidth. *ACM Transactions on Design Automation of Electronic Systems*, 8(4):399–429, October 2003.
[28] Erik Jan Marinissen. The Role of Test Protocols in Automated Test Generation for Embedded-Core-Based System ICs. *Journal of Electronic Testing: Theory and Applications*, 18(4/5):435–454, August 2002.
[29] Akihiro Ikeda et al. Design and Measurements of Test Element Group Wafer Thinned to 10μm for 3D System in Package. In *Proceedings IEEE International Conference on Microelectronic Test Structures*, pages 161–164, March 2004.
[30] Dan Perry et al. Impact of Thinning and Packaging on a Deep Sub-Micron CMOS Product. In *Electronic Workshop Digest of DATE 2009 Friday Workshop on 3D Integration*, page 282, April 2009. (http://www.date-conference.com/files/file/09-workshops/date09-3dws-digestv2-090504.pdf).
[31] Erik Jan Marinissen et al. Minimizing Pattern Count for Interconnect Test under a Ground Bounce Constraint. *IEEE Design & Test of Computers*, 20(2):8–18, March/April 2003.
[32] William H. Kautz. Testing of Faults in Wiring Networks. *IEEE Transactions on Computers*, C-23(4):358–363, April 1974.
[33] P. Goel and M.T. McMahon. Electronic Chip-in-Place Test. In *Proceedings IEEE International Test Conference (ITC)*, pages 83–90, October 1982.
[34] P.T. Wagner. Interconnect Testing with Boundary Scan. In *Proceedings IEEE International Test Conference (ITC)*, pages 52–57, October 1987.
[35] William R. Mann et al. The Leading Edge of Production Wafer Probe Test Technology. In *Proceedings IEEE International Test Conference (ITC)*, pages 1168–1195, October 2004.
[36] Brian Moore et al. High Throughput Non-Contact SIP Testing. In *Proceedings IEEE International Test Conference (ITC)*, October 2007. Paper 12.3.
[37] Erik Jan Marinissen et al. Contactless Testing: Possibility or Pipe-Dream? In *Proceedings Design, Automation, and Test in Europe (DATE)*, pages 676–671, April 2009.
[38] Dimitrios Velenis et al. Impact of 3D Design Choices on Manufacturing Cost. In *Proceedings IEEE International Conference on 3D System Integration (3DIC)*, September 2009.
[39] Xiaoxia Wu, Paul Falkenstern, and Yuan Xie. Scan Chain Design for Three-dimensional Integrated Circuits (3D ICs). In *Proceedings International Conference on Computer Design (ICCD)*, pages 208–214, October 2007.
[40] Xiaoxia Wu et al. Test-Access Mechanism Optimization for Core-Based Three-Dimensional SOCs. In *Proceedings International Conference on Computer Design (ICCD)*, pages 212–218, October 2008.
[41] Xiaoxia Wu et al. Test-Access Solutions for Three-Dimensional SOCs. In *Proceedings International Conference on Computer Design (ICCD)*, October 2008.
[42] Li Jiang, Lin Huang, and Qiang Xu. Test Architecture Design and Optimization for Three-Dimensional SoCs. In *Proceedings Design, Automation, and Test in Europe (DATE)*, pages 220–225, April 2009.
[43] Harald Vranken et al. Enhanced Reduced Pin-Count Test for Full-Scan Design. In *Proceedings IEEE International Test Conference (ITC)*, pages 738–747, November 2001.
[44] Janusz Rajski et al. Embedded Deterministic Test. *IEEE Transactions on Computer-Aided Design*, 23(5):776–792, May 2004.
[45] Yervant Zorian. A Structured Testability Approach for Multi-Chip Modules Based on BIST and Boundary-Scan. *IEEE Transactions on Components, Packaging and Manufacturing Technology – Part B*, 17(3):283–290, August 1994.
[46] Yervant Zorian and Hakim Bederr. An Effective Multi-Chip BIST Scheme. *Journal of Electronic Testing: Theory and Applications*, 10(1-2):87–95, January/February 1997.

978-1-4244-4867-8/09/$25.00 ©2009 IEEE